Vector of motion measurements in the living cochlea using a 3D OCT vibrometry system

WIHAN KIM,[1,5] DEREK LIU,[1,5] SANGMIN KIM,[2] KUMARA RATNAYAKE,[1] FRANK MACIAS-ESCRIVA,[1] SCOTT MATTISON,[2,3] JOHN S. OGHALAI,[1] AND BRIAN E. APPLEGATE[1,4,*]

[1]*Caruso Department of Otolaryngology - Head and Neck Surgery, University of Southern California, Los Angeles, CA 90033, USA*
[2]*Department of Biomedical Engineering, Texas A&M University, College Station, TX 77843, USA*
[3]*Current address: Department of Engineering and Physics, University of Central Oklahoma, Edmond, OK 73034, USA*
[4]*Department of Biomedical Engineering, University of Southern California, Los Angeles, CA. 90089, USA*
[5]*Contributed equally*
*brianapp@usc.edu

Abstract: Optical coherence tomography (OCT) has become an important tool for measuring the vibratory response of the living cochlea. It stands alone in its capacity to measure the intricate motion of the hearing organ through the surrounding otic capsule bone. Nevertheless, as an extension of phase-sensitive OCT, it is only capable of measuring motion along the optical axis. Hence, measurements are 1-D. To overcome this limitation and provide a measure of the 3-D vector of motion in the cochlea, we developed an OCT system with three sample arms in a single interferometer. Taking advantage of the long coherence length of our swept laser, we depth (frequency) encode the three channels. An algorithm to depth decode and coregister the three channels is followed by a coordinate transformation that takes the vibrational data from the experimental coordinate system to Cartesian or spherical polar coordinates. The system was validated using a piezo as a known vibrating element that could be positioned at various angles. The angular measurement on the piezo was shown to have an RMSE of $\leq 0.30°$ (5.2 mrad) with a standard deviation of the amplitude of ≤ 120 pm. Finally, we demonstrate the system for *in vivo* imaging by measuring the vector of motion over a volume image in the apex of the mouse cochlea.

1. Introduction

The inner ear plays a key role in the process of sound perception by transducing sound pressure waves into electrical signals that are passed to the brain via the auditory nerve. As sound propagates through the cochlea, passive mechanical grades tune the traveling wave with a tonotopic distribution such that high-frequency sounds peak near the base and low-frequency sounds peak near the apex [1]. An active process termed cochlear amplification further tunes the traveling wave, contributing to the frequency selectivity and high sensitivity unique to mammalian hearing [2,3]. However, the exact mechanisms of cochlear amplification remain unclear, in large part due to the limitations of current techniques used to measure vibrations within the cochlea [4].

Early studies of the cochlea exploited Laser Doppler Vibrometry (LDV), which required opening the cochlea to place a reflective bead on the structure of interest [5–7]. However, this invasive technique altered intracochlear pressures and changed the physiological response to acoustic stimuli. Further, LDV is only capable of measurement from highly reflective structures, most commonly a glass bead placed on the surface of the structure of interest. This limits its ability to characterize structures deeper within the Organ of Corti (OoC). More recently, the development of optical coherence tomography (OCT) and vibrometry has overcome some of

these challenges [8,9]. This technique measures vibration without opening the cochlea, often through the otic capsule bone or the round window membrane. Importantly, OCT is capable of depth-resolved measurements of morphology and function through the entire thickness of the OoC.

Both LDV and OCT, however, are only capable of measuring displacement along the optical axis. While previous studies largely focused on the transverse motion of the basilar membrane (BM), recent studies using OCT showed that the structures within the OoC have complex vibratory characteristics [10–12]. These vibration patterns likely involve three-dimensional components in transverse, radial, and longitudinal directions, and their study is essential to elucidating the mechanisms of cochlear amplification. For example, inner hair cells (IHCs), the primary auditory afferent cells, are stimulated by radial deflection of their stereocilia. Previous attempts to characterize three-dimensional motion included one study utilizing LDV, although its use of opened *ex vivo* preparations limits interpretation [13]. Utilizing OCT, vibration characteristics have been studied in each direction, although one-dimensional measurements are prone to variations in anatomical orientation relative to the optical axis, and attempts to reconstruct two- or three-dimensional motion require separate, repeated measurements [14,15]. The experimental complexity of serial multi-dimensional measurements erects a practical barrier to executing studies that account for the entire vector of motion in the OoC rather than a single component of the vector. This motivated the work presented here, where we transpose the complexity onto the optical system such that 3-D measures of the OoC vector of motion can become routine. The aspiration of this work was to make the 3-D measurements as routine as the current 1-D measurements made daily in our lab.

While we are not aware of other work to develop 3-D OCT vibrometry, there are other related works similarly built on phase-sensitive OCT that try to resolve multiple axes in a single measurement. Trasischker et al. and Haindl et al. implemented a three-channel Doppler OCT system by integrating three separate spectral domain OCT systems [16,17]. This allowed them to resolve the flow velocity vector in flow phantoms and the retina of a human volunteer, but at the cost of system complexity with three independent spectrometer detectors and light sources. A depth-encoded method was demonstrated by Ahn et al, using a single OCT system with a beam divider in the sample arm that generated 5 viewpoints at 5 different depths [18]. The relatively small depth range of the spectral domain system, 3.4 mm, which had to be split into 5 components, strongly limited thickness of samples that could be probed. Finally, a time-encoded method used optical switches in swept-source OCT system to switch between three channels. This avoided reductions in optical power and achievable imaging depth at the expense of prolonged measurement times [19]. As a result, this method was vulnerable to motion artifact because of the 3-fold reduction in sweep rate.

Here we describe our approach to a similar problem. Since we are measuring vibrations in the cochlea of the mouse, not flow, our sweep rate has to be at least 10's of kilohertz with a usable depth of at least a couple of millimeters. To fulfill these requirements, we developed a depth-encoded OCT vibrometry system using three independent optical axes in the sample arm but within a single interferometer. The system employs one swept-laser that sweeps linearly in wavenumber, has high optical power, and has a long coherence length. The high optical power compensates for reduced power directed to each of three channels while the long coherence length allows imaging over a large total depth without losses due to signal roll-off. The developed system was validated with a piezo electric element and used to measure the three-dimensional vibrations within the living mouse organ of Corti.

2.   Methods

2.1.   *Optical configuration and sound generation*

The optical system schematic for the 3D vibrometry system is shown in Fig. 1. The light source was a swept laser (Insight Photonic Solutions Inc.) centered at 1310 nm with a bandwidth of 90 nm, a stable linear sweep, and adjustable sweep rate. It has a fixed sampling rate of 400 MHz, hence adjusting the sweep rate also changes the sampling in k, δk. For this application, we adjusted the sweep rate to 41 kHz resulting in a maximum depth of 4.3 cm ($\delta k = 5.83\,m^{-1}$). This was necessary to allow enough z-space to incorporate three sample arms (channels) into one interferometer (explanation below). The laser output was launched into a 1×2, 90:10 fiber coupler with 10% of the output directed to the reference arm. The reference arm incorporated a circulator and free-space portion that was designed to nearly match the dispersion of the sample arm optical system. The reference arm light was recombined with the sample arm at a 4×4 fiber coupler.
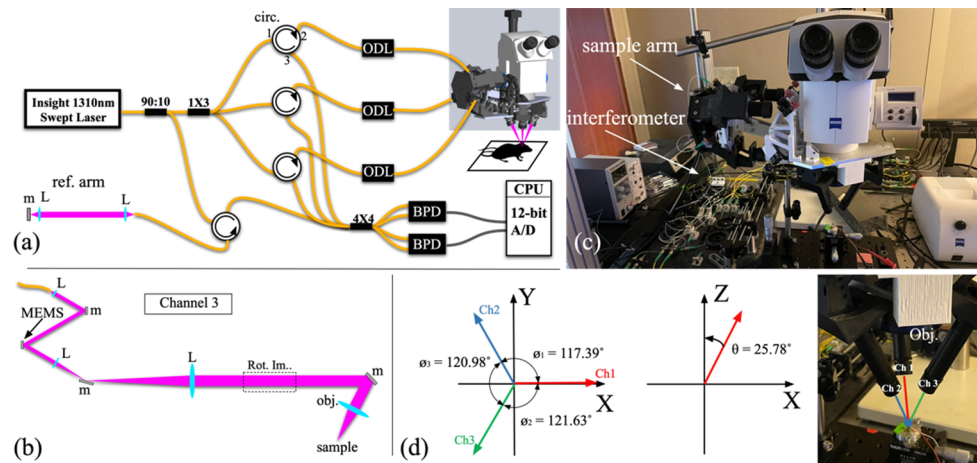

Fig. 1. (a) Optical schematic of 3D OCT vibrometry system. Abbreviations: circ, optical circulator; ODL, optical delay line; m, mirror; L, lens; BPD, balanced photodetector (b) Optical schematic for Channel 3 sample arm. (c) Photograph of system attached to Zeiss stereo microscope. Arrows indicate the location of fiber optic portion of *interferometer* and free-space portion of *sample arm*. (d) Relative orientation of 3 sample arms where z corresponds to the optical axis of the stereo microscope. Labeled image shows position of 3 sample arm objectives (Ch1, Ch2, Ch3) and stereo microscope objective.

Ninety percent of the laser light was directed to the sample arm, entering a 1×3 fiber coupler. The 3 output ports of the coupler illuminate 3 separate sample arms. Each sample arm had a circulator followed by polarization control (not shown) and a fiber based optical delay line (ODL). The ODLs allow the relative depth of the focal plane in each channel to be independently and precisely chosen so that images from each do not overlap. The output of the delay line entered the free-space portion of the sample arm as shown in Fig. 1(b). This schematic specifically shows the beam path for Channel 3, however all three channels had similar optics. Light exiting the fiber was collimated. A folding mirror directed the light onto a MEMS mirror used to scan the beam across the sample. All three sample arms use the same MEMS mirror. This ensures perfectly synchronous and repeatable imaging. After exiting the MEMS mirror, the light enters the first lens in a beam expander with a magnification of 3. The second lens recollimates the light. In Channel 3 only, there is a set of mirrors that rotates the image (Rot. Im.). These were inserted to enable all three channels to scan along a single line. This was done to allow vibrometry along a

segment of the Organ of Corti without recording an entire volume (full explanation below). The objectives for the three sample arms are arranged around the objective of a Zeiss Discovery V20 stereo microscope as shown in the inset of Fig. 1(b). The optical axis of the three sample arms intersect at their focal plane and are designed to have a 120° separation in the horizontal plane and 27° relative to the optical axis of the Zeiss microscope as shown in Fig. 1(c). The focal spot of each was 15 ± 1 µm (FWHM) with a corresponding depth of focus of ∼800 µm. The standard deviation of the phase noise measured in each channel was within a factor of ∼2 of the theoretical limit, similar to what we have measured before [20,21]. The Zeiss objective was removed from the body of the microscope and placed a few centimeters lower to allow the light from Channel 3 to pass through the microscope. The only impact on the image through Zeiss microscope was a small decrease in the field of view.

Back-reflected light from the sample passes back through the sample arm optics, enters the circulator, and exits through port 3. The three sample arms enter the remaining 3 ports of the 4×4 coupler where they are combined with the reference arm. The output of the 4×4 fiber coupler is converted to electrical signals through two balanced detectors (Weiser Lab, BPD600MA). The electrical signals are digitized at 400 MS/s by an Alazartech card (ATS9373, Alazartech Inc.) and processed on a graphic processing unit (GPU). The whole system runs on custom software written in Python, with time-critical and computationally-intensive processes implemented in C and CUDA C.

The relative angles of the 3 channels were determined in the following way. First, we measured the beam paths by placing a camera on an optical stage at the intersection of the three beams and translating the camera vertically, recording an image at each position. The vertical position of the camera was precisely controlled using a micrometer with 127 µm between vertical steps for a total translation of 1.8 mm over 15 steps. This produced an image stack of the beam geometry that could be analyzed to find the relevant angles. The analysis script was written in MATLAB. For each image the centroid of each of the three spots were found to generate a set of 3-D coordinates for samples along the three beams. The samples were grouped into 3 sets, one for each channel. Samples from adjacent channels were fit to a plane to yield a set of 3 planes whose intersections define the optical axes of the three channels. The intersections were found and used to derive 3-D equations for the optical axes. Once these were found, it was simple to use the law of cosines to find the angles shown in Fig. 1(d). This measurement of the system angles needs to be redone each time the optical system is realigned. It typically changes by less the 2 degrees.

2.2. *Signal-processing*

The combination of the three channels in a single interferometer necessitates a few additional calibration/processing steps compared to our typical 1D vibrometry system. Figure 2 shows a flow chart depicting the processing steps. The first three operations shown on the flow chart are common to any spectral domain OCT system (see for instance [22] and references therein), i.e. interferograms, H(x,y,k,t), are background subtracted, dispersion compensated, and inverse fast-Fourier transformed (iF_k) to transform from the spatial frequency domain (k-domain) signals to the space domain (z-domain). The dispersion for each channel is slightly different so the dispersion compensation and iF_k is done three times, one for each channel to yield $h_n(x_n,y_n,z_n,t)$, where n is the channel. These complex signals are decoded or divided spatially by depth into three channels.

Each channel in the interferometer has a different optical delay set by the ODL, so that three images appear in a B-scan, one from each channel. From an image with 7000 samples in depth, indices were predetermined such that each channel consisted of 1001 samples, corresponding to a total of approximately 3 mm, and separated from each consecutive channel by 1000 additional samples. These values were chosen to maintain sufficient depth of each channel as well as sufficient separation between channels. The ODLs were individually adjusted such that the
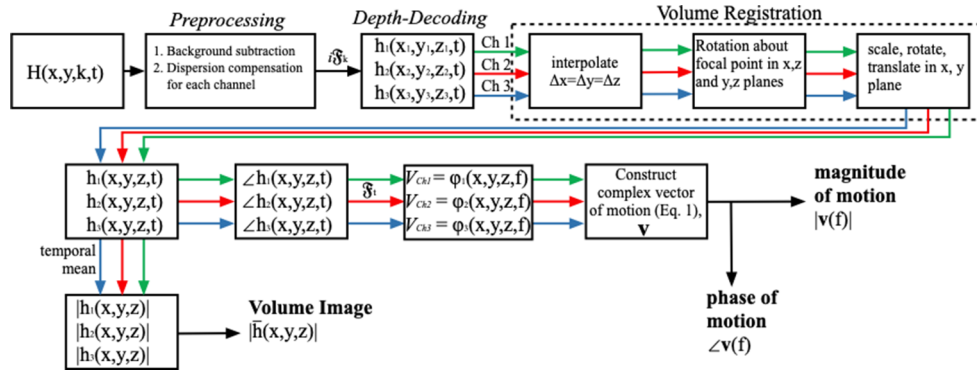
Fig. 2. Processing flow chart of volumetric M-scan data, H(x,y,k,t). iF_k, inverse Fourier transform along the k-dimension. F_t, Fourier transform along the time dimension.

image of the reflector was centered in the chosen depth for each channel. Once the channels are separated out in depth, they are processed separately and then combined to generate an average volume image ($|\bar{h}|$) and the complex vector of motion (**v**) following the processing flow chart in Fig. 2 where the channels are color coded to match Fig. 1(d).

After obtaining $h_n(x_n,y_n,z_n,t)$ we co-register the three volume M-scans (dashed box Fig. 2). First, the data are interpolated using bicubic methods to have symmetric voxels, such that the sampling in all dimensions is the same, i.e. $\Delta x = \Delta y = \Delta z$. It is then rotated on the xz- and yz-planes to compensate for the tilt due to the beam geometry in Fig. 1. The rotated data are registered with pre-calculated image registration coefficients on the xy-plane to compensate for rotation caused by the beam geometry and distortion caused by scanning. A detailed explanation of the calibration that generates the rotation angles and image registration coefficients is provided below. At this point the coordinate systems and sampling of the three channels are identical, hence we drop the subscripts on x,y,z, e.g. $h_n(x,y,z,t)$. The volume image is created by first taking the mean along time ($|h_n(x,y,z)|$) followed by averaging the magnitude of the three co-registered volumes.

The phase of the co-registered M-scans, $\angle h_n(x, y, z, t)$, is then unwrapped and Fourier transformed along the time-dimension as previously described [8]. We then have the frequency dependent vibrational amplitude and phase, V_n ($n = ch1, ch2$, ch3), for each channel at every voxel in the volume image. To construct a complex vector of motion at each voxel, we must do a coordinate transformation that takes us from the experimental coordinate system defined by the three channels (Fig. 1(c) and 3(a)) into Cartesian or spherical polar space. Equation (1a) was written by inspection and relates the experimental coordinate system to Cartesian space-fixed axis system where the transformation is defined at the voxel level. In other words, the array on the left are the vibrational amplitudes and phases from the three channels at a single spatial sample in the volume. Performing a matrix inversion or solving the system of equations implied by eq 1a, yields Eq. (1b) which will allow us to directly transform the vibrational motion at each voxel into the Cartesian fixed-space axis system where the z-axis is referenced to the optical axis of the stereomicroscope. The resulting Cartesian vector can be transformed into its spherical polar equivalent using the relations in Eq. (1c) (Fig. 3(c)) reproduced here to make it clear what convention we are using. Finally, we have the vector of motion, **v** (Eq. (1d)). Recall **v** corresponds

to periodic motion, hence it has a magnitude and phase.

$$\begin{pmatrix} V_{Ch1} \\ V_{Ch2} \\ V_{Ch3} \end{pmatrix} = \begin{pmatrix} \sin\theta & 0 & \cos\theta \\ \sin\theta\cos\phi_2 & \sin\theta\sin\phi_2 & \cos\theta \\ \sin\theta\cos\phi_3 & \sin\theta\sin\phi_3 & \cos\theta \end{pmatrix} \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} \tag{1a}$$

$$\begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} = \frac{1}{(\sin(\phi_2 - \phi_3) - \sin(\phi_2) + \sin(\phi_3))} \begin{pmatrix} \frac{-\sin(\phi_2)+\sin(\phi_3)}{\sin(\theta)} & \frac{-\sin(\phi_3)}{\sin(\theta)} & \frac{\sin(\phi_2)}{\sin(\theta)} \\ \frac{\cos(\phi_2)-\cos(\phi_3)}{\cos(\theta)} & \frac{\cos(\phi_3)-1}{\sin(\theta)} & \frac{1-\cos(\phi_2)}{\sin(\theta)} \\ \frac{\sin(\phi_2-(\phi_3))}{\cos(\theta)} & \frac{\sin(\phi_3)}{\cos(\theta)} & \frac{-\sin(\phi_2)}{\cos(\theta)} \end{pmatrix} \begin{pmatrix} V_{ch1} \\ V_{ch2} \\ V_{ch3} \end{pmatrix} \tag{1b}$$

$$V_r = \sqrt{V_x^2 + V_y^2 + V_z^2}, V_r \in (0, \infty)$$

$$V_\theta = \cos^{-1}\left(\frac{V_z}{r}\right), \quad V_\theta \in (0, \pi)$$

$$V_\phi = \begin{cases} \tan^{-1}\left(\frac{V_y}{V_z}\right) & V_x > 0 \\ \tan^{-1}\left(\frac{V_y}{V_z}\right) + \pi & V_x < 0 \\ \frac{\pi}{2} & V_x = 0 \end{cases}, V_\phi \in (0, 2\pi) \tag{1c}$$

$$\mathbf{v} = V_x\hat{\mathbf{x}} + V_y\hat{\mathbf{y}} + V_z\hat{\mathbf{z}} = V_r\hat{\mathbf{r}} + V_\theta\hat{\boldsymbol{\theta}} + V_\phi\hat{\boldsymbol{\phi}} \tag{1d}$$
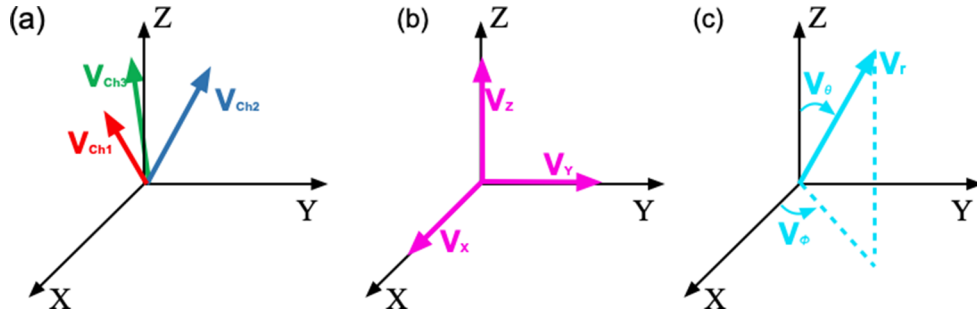

Fig. 3. Coordinate systems referenced in Eq. (1) (a) Experimental (b) Cartesian (c) Spherical Polar.

2.3. Volume registration

Structural and vibrational images from three separate channels were registered based on predetermined values for depth decoding, tilt angles, and an affine transformation as shown schematically in Fig. 4(a). Tilt angles were first determined using the central B-scan in the x,z-plane, Fig. 4 (a). The angle of the reflector relative to the x-axis was used to perform the first rotation. Note, the sample was placed in the focal plan of the instrument and all rotations were done about the central point on the focal plane. The process was repeated with the central B-scan in the y,z-plane to perform the second rotation (not shown). The coefficients of the affine transformation were determined using a paper United States Air Force Resolution Test Chart to provide easily identified features in all three channels. A minimum of six matching control points in each channel were manually selected and used to infer the parameters of the affine transformation, which is applied to channel 1 and channel 3 to match a reference, channel 2.
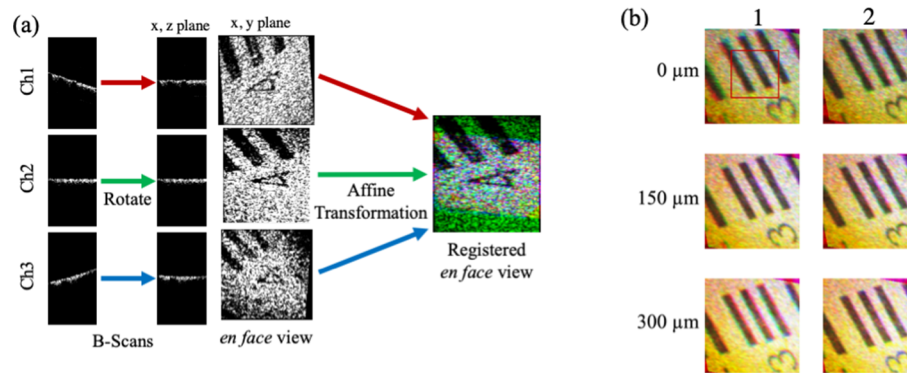
Fig. 4. Volume Registration Process (a) Schematic showing steps taken to calibrate using a US Air Force resolution target. These steps are described with more detail in the text. (b) Images of an air force test chart extracted from volume images where the extracted image is in the x,y plane. The test chart was translated along the z-axis in 150 μm increments. Column 1, images registered using a single pair affine transformations for all volume sets. Column 2, images registered using affine transformation pairs optimized for each volume set. Images are 1.5×1.5 mm^2.

In an ideal optical system there are no aberrations and we would expect the affine transformation to work similarly at every depth across the entire field of view. In order to check for imperfections in the registration due to optical aberration, we translated the Test Chart along the z-axis, collecting volume images at multiple positions separated by 150 μm. Representative cuts in the x,y plane from three volumes are shown as RGB images in Fig. 1(b). Each volume was smoothed using a $3 \times 3 \times 5$ voxel kernel ($16.3 \times 16.3 \times 13.9$ μm^3) to improve signal to noise. The first column used a single pair of affine transformations (as above) to register the volume images from the three channels at each depth. The second column used a pair of affine transformations calculated from each volume set. In other words, the images in each row of column 1 were registered with the same pair of affine transformations, while the images in each row of column 2 were registered with a unique pair of affine transformations optimized for that volume image. Column 2 represents the best we could hope to do with an affine transformation. In the absence of aberration these two rows should be identical within the error introduced by hand selecting control points. The entire field of view is 1.5×1.5 mm^2, however we only used the central 25% of this for imaging the cochlea (indicated by the red box in the first image). If we look carefully at the long edges of the bars at the center of the images within the area that would be encompassed by this box, we can see a maximum deviation of ~23 μm in column 1 (300 μm image) and ~20 μm in column 2 (0 μm image). This is a factor of 1.3-1.5 of the lateral optical resolution over a depth range of 300 μm. Stronger aberration is expected at the periphery of the lens, hence we expect worse performance outside the central 25%. This is clearly most evident in column 1 in several places, e.g. upper left of the 0 μm image. Nevertheless performance is reasonable within the central area of the field of view used for mouse cochlear imaging over a depth larger than the thickness of the organ of Corti.

Next, we compared the registration performance on a mouse cochlea, observing how the three volumes overlapped. A representative frame is shown in Fig. S1(a) for all three channels. The overlap is shown by the RGB overlay in Fig. S1(b). The regions in the RGB overlay which are not white appear to be largely due to differences in image intensity rather than poor overlap. The light illuminating the sample from each channel enters the otic capsule at different angles, hence must travel through different thicknesses of bone to reach the organ of Corti. This will undoubtedly result in different signal intensities in the three channels. We have attempted to illustrate this by
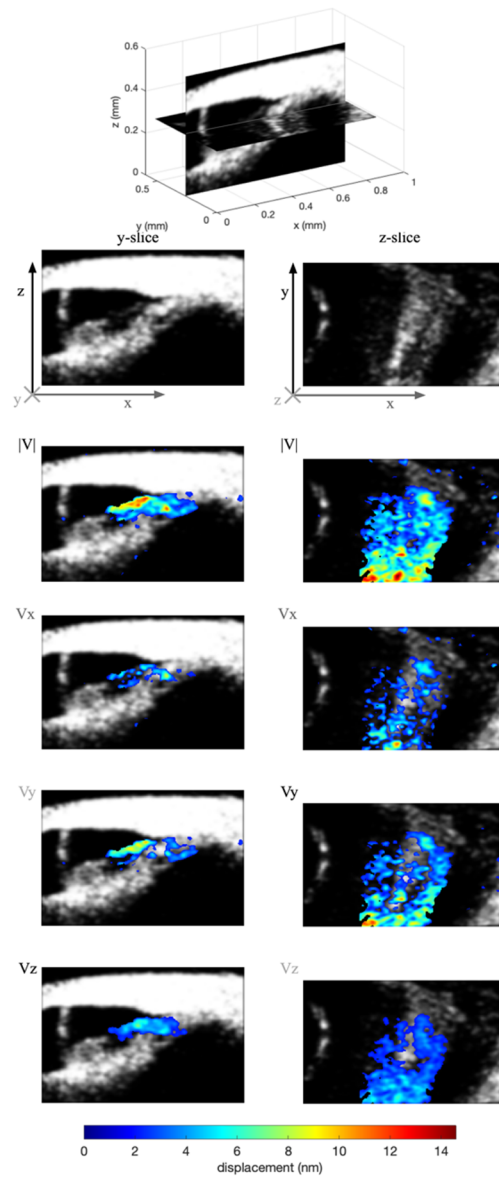
Fig. 5. In vivo mouse cochlea images and components of vector of motion measured with a 9 kHz, 100 dB SPL sound stimulus. Gray scale are OCT images. Color scale are vibrational amplitudes.

rendering an RGB overlay of binary masks created from each channel, Fig. S1(c). In the binary mask there is good overlap in the organ of Corti. The tectorial membrane has some overlap but not as complete as the organ of Corti. Finally, given the difficulty in making quantitative comparisons on individual cross-sections we rendered the three registered volumes using the Amira software package and visually inspected for overlap. In the volume data it is easier to observe the overlap on larger readily identifiable features. Overall, we considered the agreement between the volumes to be good where we checked which gave us confidence that the registration process was acceptable for the mouse cochlea.

3. Results and discussion

3.1. Phantom experiment

To verify the system performance and estimated accuracy, we used a piezo mounted on a goniometer and rotation stage to mimic a vibrating sample at various angles. A paper US Air Force test chart was attached to the surface of the piezo as a scattering target. For this experiment, the piezo was driven by a sinusoidal signal with the frequency of 5 kHz to generate a periodic vibration. A volume M-scan was collected using the same parameters and dimensions as the volume image used for calibration above. The registration coefficients generated during the calibration process were used to register the volume M-scan. Following registration, a central square region of interest approximately 250 microns in length and width was chosen. For each channel the median amplitude was calculated within this region of interest. These three values were then used to calculate single vectors of motion using Eq. (1). Angles between vectors were calculated using the law of cosines.

The piezo was placed on a dual-axis goniometer (GNL20, Thorlabs) to tilt the piezo to a known polar angle (θ), and on a rotation stage (RP01, Thorlabs) to rotate the piezo to a known azimuthal angle (ϕ). First, the polar angle was changed by varying the goniometer along a single axis from 0 to 10 degrees in increments of 2.5 degrees. Varying the azimuthal angle with a polar angle of 0 degrees would result in no measurable change, thus the azimuthal angle was changed by varying the rotation stage from 0 to 60 degrees in increments of 15 degrees while holding the goniometer at its maximum of 10 degrees.

It is not possible to perfectly place the piezo element precisely parallel to the optical axis of the microscope (Z-axis, Fig. 3), hence 0 in Table 1 refers to 0 on the goniometer and rotation stage. The measurement at 0 was used as a reference, so all reported angles are relative to this first measurement. The angle between the two vectors were calculated with the law of cosines. This result is reported as *Measured Relative Angle* in Table 1. The *Expected Relative Angle* is the same calculation but based upon the known rotation of the goniometer and rotation stage. The final column in Table 1 is the absolute error. Based on these, the RMSE was 0.24° (4.2 mrad) when changing the polar angle and 0.30° (5.2 mrad) when changing the azimuthal angle. Our ability to position both the goniometer and rotation stage was likely no better than ±0.2°, hence we consider this result to be good and certainly acceptable for our application.

We also calculated the amplitude of vibration for these measurements. These values are given in the first column of Table 1. For the experiments varying the polar and azimuthal angles the mean and standard deviation was 24.15 ± 0.12 nm and 23.66 ± 0.06 nm, respectively. Nominally, since the same voltage was applied to the piezo for both experiments the displacement should be the same, however since we conducted these experiments at different times small changes in environmental factors such as temperature likely led to the small shift. We also expect the change in angle to slightly change the load on the piezo due to gravity, which also leads to some experimental error. Again, we consider this to be a good result and adequate for our application.

Table 1. Measured change in vector angles using a piezoelectric actuator.

	Amplitude, r (nm)	Expected Relative Angle (deg)	Measured Relative Angle (deg)	Absolute Error (deg)
Polar, θ	24.06	0	Reference	Reference
	24.10	2.50	2.71	0.21
	24.09	5.00	5.14	0.14
	24.36	7.50	7.55	0.05
	24.16	10.00	9.59	-0.41
Azimuthal, ϕ	23.63	0	Reference	Reference
	23.57	1.82	1.86	0.04
	23.73	3.60	3.15	-0.45
	23.68	5.32	5.17	-0.15
	23.70	6.96	6.58	-0.38

3.2. Animal experiment

The 3D OCT vibrometry system was then used to measure motion in a mouse cochlea. All study protocols were approved by the University of Southern California Institutional Animal Care and Use Committee. The mouse cochlea was accessed using a previously reported approach with one key modification to account for the 3D OCT system design. In brief, a P28-P42 CBA/CaJ wild-type mouse was anesthetized with a ketamine/xylazine mixture. Supplemental doses were administered to maintain areflexia. The mouse was secured in a custom head holder, and a ventrolateral surgical approach was used to expose the middle ear bulla. In previous studies, the bulla was opened medial to the tympanic annulus, which maintained the integrity and attachment of the tympanic membrane. This window was sufficient for the passage of a single OCT beam aligned with the optical axis, but not for the simultaneous passage of three OCT beams positioned at different polar and azimuthal angles. Thus, the opening of the bulla was carefully extended laterally to include portions of the tympanic annulus and bony external auditory canal, resulting in a minor conductive hearing loss. This loss was found to be ~10-15 dB SPL from comparisons of tuning curves before and after extending the opening laterally. Digitally generated sound stimuli were presented via a speaker (FT17H, Fostex) connected with plastic tubing to an acoustic probe containing a microphone (ER10B+, Etymotic). The probe was sealed to the residual ear canal with dental cement. Sound stimuli were calibrated using the pressure measured by the probe's microphone after correcting for its frequency-dependent sensitivity [23]. A volume M-scan was measured while presenting a 100 dB SPL sound stimulus at 9 kHz, which is the characteristic frequency of the cochlear region of study. The 3D images and vibration vectors were registered and transformed as described above. The mouse was euthanized by anesthetic overdose, and a set of post-mortem volume M-scans were collected.

The top image in Fig. 5 is a pair of cross-sectional OCT images taken from a volume M-scan processed as shown in Fig. 2, i.e. $|\bar{h}(x, y, z)|$. The 3-D rendering indicates the relative position of the two slices which also shown just below as 2-D gray scale images. The labeled axes on the 2-D images indicate relative orientation where the axis system is relative to the optical axis of the stereomicroscope where z lies on the optical axis. The speckle is noticeably reduced compared to normal expectations for a typical OCT image. The volume images from the three channels have different speckle patterns since they are acquired at different angles. Therefore, the averaging that takes place in the final step to produce $|\bar{h}(x, y, z)|$ reduces speckle. The final processing step (not shown in Fig. 2) was a 3-D convolution with a symmetric Gaussian kernel having a standard deviation of 6.4 μm and 22.4 μm length along each dimension. This was done

to take advantage of our spatial oversampling to improve SNR with only a modest decrease in resolution to 16 μm from 15 μm.

Just below in Fig. 5 is an overlay of the vibrational magnitude, thresholded at 2 nm, onto the corresponding OCT image. This threshold corresponds to the upper limit of motion measured in the bone of the otic capsule as is evident in Fig. S2. The individual vector components are shown below, using the same threshold. These were generated from the vibrational amplitudes of the three channels using Eq. (1b) followed by the same Gaussian convolution described above. The colorbar at the bottom of the figure applies to all vibrational images. If we compare the anatomical features to our prior work (e.g. Figure 3(E)-(F) in Ref. [10]), the images recorded with our 1-D OCTV system appears to correspond most closely to the V_z component, although it clearly also includes a significant contribution from V_x. Since the overlay images inherently loose information due to the threshold we also provide images of the vibrational amplitude and phase as Figs. S2 and S3. Comparisons of the vector components demonstrate clear differences in the measured motion along the three orthogonal directions. One obvious feature from our 1-D measurements that is reproduced in the phase of V_x and V_z is the large phase shift at the edge of the organ of Corti (magenta in V_x and dark blue in V_z) compared to the Basilar membrane. The data shown is for a single mouse at one frequency and sound pressure level. We have collected a much larger set of data on multiple mice which will be published elsewhere, focusing on the contribution to our understanding of cochlear mechanics.

4. Conclusions

The 3-D OCTV system we have developed is capable of spatially resolving the nanometer scale motion within the living murine cochlea enabling the construction of a vector of motion. This improved capacity for measuring cochlear vibrations comes at the cost of a significantly more complicated optical system. However, it enables much easier experimental measurement of the vector of motion and likely more reliable, than acquiring multiple 1-D measurements at multiple angles as has been done before [14].

The approach takes advantage of the long coherence length and tunable sweep rate of the Insight laser to depth encode three channels, where the channels have been arranged to optically probe the sample at three different angles. In the current setup we were very conservative in our spacing of the three channels in depth. As a consequence, the laser sweep rate was set to 41 kHz resulting in a maximum vibrational frequency of 20.5 kHz. This is sufficient for imaging the apex of the mouse cochlea where we typically work. However, for imaging closer to the base we would need a larger maximum frequency. It is likely we could improve this by a factor of 2 within the current system to achieve a maximum vibrational frequency of 41 kHz without any significant reduction in performance. In order to reach frequencies up to 90 kHz as could be measured at the base, we would need a laser with approximately twice the sampling rate, i.e. 800 MHz.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not currently publicly available but may be obtained from the authors upon reasonable request

Supplemental document. See Supplement 1 for supporting content.

References
1. G. Von Békésy, *Experiments in Hearing* (McGraw Hill, 1960).
2. J. Ashmore, P. Avan, W.E. Brownell, P. Dallos, K. Dierkes, R. Fettiplace, K. Grosh, C.M. Hackney, A.J. Hudspeth, F. Jülicher, B. Lindner, P. Martin, J. Meaud, C. Petit, J. Santos-Sacchi, J.R. Santos Sacchi, and B. Canlon, "The remarkable cochlear amplifier," Hear. Res. **266**(1-2), 1–17 (2010).

3. J.S. Oghalai, "The cochlear amplifier: augmentation of the traveling wave within the inner ear," Curr Opin Otolaryngol Head Neck Surg **12**(5), 431–438 (2004).

4. J.J. Guinan, "The interplay of organ-of-Corti vibrational modes, not tectorial- membrane resonance, sets outer-hair-cell stereocilia phase to produce cochlear amplification," Hear. Res. **395**, 108040 (2020).

5. S.S. Narayan, A.N. Temchin, A. Recio, and M.A. Ruggero, "Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae," Science **282**(5395), 1882–1884 (1998).

6. W. He, A. Fridberger, E. Porsove, K. Grosh, and T. Ren, "Reverse wave propagation in the cochlea," Proc. Natl. Acad. Sci. U. S. A. **105**(7), 2729–2733 (2008).

7. T. Ren, W. He, and P.G. Gillespie, "Measurement of cochlear power gain in the sensitive gerbil ear," Nat. Commun. **2**(1), 216 (2011).

8. S.S. Gao, P.D. Raphael, R. Wang, J. Park, A. Xia, B.E. Applegate, and J.S. Oghalai, "In vivo vibrometry inside the apex of the mouse cochlea using spectral domain optical coherence tomography," Biomed. Opt. Express **4**(2), 230–240 (2013).

9. S.S. Gao, R. Wang, P.D. Raphael, Y. Moayedi, A.K. Groves, J. Zuo, B.E. Applegate, and J.S. Oghalai, "Vibration of the organ of Corti within the cochlear apex in mice," J. Neurophysiol. **112**(5), 1192 (2014).

10. H.Y. Lee, P.D. Raphael, J. Park, A.K. Ellerbee, B.E. Applegate, and J.S. Oghalai, "Noninvasive in vivo imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea," Proc. Natl. Acad. Sci. U. S. A. **112**(10), 3128–3133 (2015).

11. T. Ren, W. He, and D. Kemp, "Reticular lamina and basilar membrane vibrations in living mouse cochleae," Proc. Natl. Acad. Sci. U. S. A. **113**(35), 9910–9915 (2016).

12. N.P. Cooper, A. Vavakou, and M. van der Heijden, "Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea," Nat. Commun. **9**(1), 3054 (2018).

13. W. Hemmert, H.P. Zenner, and A.W. Gummer, "Three-dimensional motion of the organ of Corti," Biophys. J. **78**(5), 2285–2297 (2000).

14. H.Y. Lee, P.D. Raphael, A. Xia, J. Kim, N. Grillet, B.E. Applegate, and J.S. Oghalai, "Two-dimensional cochlear micromechanics measured in vivo demonstrate radial tuning within the mouse organ of corti," J. Neurosci. **36**(31), 8160–8173 (2016).

15. J.B. Dewey, A. Xia, U. Müller, I.A. Belyantseva, B.E. Applegate, and J.S. Oghalai, "Mammalian auditory hair cell bundle stiffness affects frequency tuning by increasing coupling along the length of the cochlea," Cell Rep. **23**(10), 2915–2927 (2018).

16. R. Haindl, W. Trasischker, A. Wartak, B. Baumann, M. Pircher, and C.K. Hitzenberger, "Total retinal blood flow measurement by three beam Doppler optical coherence tomography," Biomed. Opt. Express **7**(2), 287–301 (2016).

17. W. Trasischker, R.M. Werkmesiter, S. Zotter, B. Baumann, T. Torzicky, M. Pircher, and C.K. Hitzenberger, "In vitro and in vivo three-dimensional velocity vector measurement by three-beam spectral-domain Doppler optical coherence tomography," J. Biomed. Opt. **18**(11), 116010 (2013).

18. Y.C. Ahn, W. Jung, and Z. Chen, "Quantification of a three-dimensional velocity vector using spectral-domain Doppler optical coherence tomography," Opt. Lett. **32**(11), 1587–1589 (2007).

19. A. Wartak, F. Beer, B. Baumann, M. Pircher, and C.K. Hitzenberger, "Adaptable switching schemes for time-encoded multichannel optical coherence tomography," J. Biomed. Opt. **23**(05), 1–12 (2018).

20. J. Park, E.F. Carbajal, X. Chen, J.S. Oghalai, and B.E. Applegate, "Phase-sensitive optical coherence tomography using an Vernier-tuned distributed Bragg reflector swept laser in the mouse middle ear," Opt. Lett. **39**(21), 6233–6236 (2014).

21. S. Kim, J.S. Oghalai, and B.E. Applegate, "Noise and sensitivity in optical coherence tomography based vibrometry," Opt. Express **27**(23), 33333–33350 (2019).

22. W. Drexler and J.G. Fujimoto, "Optical coherence tomography technology and applications," in *Biological and Medical Physics, Biomedical Engineering* (Springer, 2008).

23. D.M. Rasetshwane and S.T. Neely, "Calibration of otoacoustic emission probe microphones," J. Acoust. Soc. Am. **130**(4), EL238–EL243 (2011).